From Flexible to Stiff: Systematic Analysis of Structural Phases for Single Semiflexible Polymers

Daniel T. Seaton,[1,*] Stefan Schnabel,[2,†] David P. Landau,[3,‡] and Michael Bachmann[3,§]

[1]*Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[2]*Institut für Theoretische Physik, Universität Leipzig, Postfach 100920, D-04009 Leipzig, Germany*
[3]*Center for Simulational Physics, The University of Georgia, Athens, Georgia 30602, USA*

(Received 19 August 2012; published 10 January 2013)

Inspired by recent studies revealing unexpected pliability of semiflexible biomolecules like RNA and DNA, we systematically investigate the range of structural phases by means of a simple generic polymer model. Using a two-dimensional variant of Wang-Landau sampling to explore the conformational space in energy and stiffness within a single simulation, we identify the entire diversity of structures existing from the well-studied limit of flexible polymers to that of wormlike chains. We also discuss, in detail, the influence of finite-size effects in the formation of crystalline structures that are virtually inaccessible via conventional computational approaches.

DOI: 10.1103/PhysRevLett.110.028103 PACS numbers: 87.15.−v, 02.70.Uu, 05.70.Fh, 64.60.Cn

The understanding of biomolecular function is inevitably based on the detailed knowledge of the mechanical response of biomolecular systems to thermal fluctuations. The competition between entropic and energetic system properties establishes structural phases. Their stability and associated transition barriers between them determine the ability of the system to respond to changes in the external environment. Given the complexity of the chemical structure of biomolecules, the identification and classification of the structural phases by means of thermodynamic analysis is difficult. The highly debated problem of protein folding is the most prominent example in this regard, even more so as even complex protein models often fail to make correct structure predictions. It is, therefore, rather surprising that several dynamic and structural properties of nucleic acids such as DNA and RNA have been successfully described by a comparatively simple approach, the Kratky-Porod or wormlike-chain (WLC) model [1]. In this model, the energy is determined by thermally excited bending only, $E = \frac{1}{2} k_B T \bar{\kappa} \int_0^L ds (\partial_s \mathbf{u})^2$, where T is the temperature, $\bar{\kappa}$ is the bending stiffness, and L is the contour length of the polymer chain, which is represented by the continuous curve $\mathbf{x}(s)$ with unit tangent vector $\mathbf{u}(s) = \partial_s \mathbf{x}(s)$. The correlation function $\langle \mathbf{u}(s)\mathbf{u}(s + \Delta s)\rangle$ decays exponentially with Δs, defining a characteristic correlation length scale ξ, called the persistence length. In three-dimensional embedding space, $\xi = \bar{\kappa}$. Since $\bar{\kappa}$ is a material constant, this result thus also implies that $\xi = \text{const}$; i.e., it is independent of L (and thus the number of nucleotides N) and T. This appears to be an oversimplified result, but the structural behavior of the most famous example, long [$\mathcal{O}(10^3)$ base pairs] double-stranded DNA with $\xi \approx 50$ nm, seems to be well described under this assumption.

Bending-energy-dominated polymers are called semiflexible polymers. DNA, RNA, and even some protein complexes, such as myosin fibers and actin filaments, belong to this category. However, lacking volume exclusion and potential nonbonded monomer-monomer interactions, the WLC model is generally unable to describe structural transitions of single polymers, e.g., the structural phase space of interacting, flexible polymers. This is a striking problem and motivation for our study, as recent studies of single- and double-stranded nucleic acids reveal that persistence lengths might not be constant but depend on N and external conditions such as temperature and salt concentrations [2], implying a greater flexibility than previously thought. Finite-size effects induced by small chain lengths and the actual discrete nature of polymer chains (allowing for ''kinks'' and cyclization in DNA [3–6]), as well as excluded-volume effects causing an effective thickness due to extended side chains (e.g., bottle-brush polymers [7] and proteins [2]) are also potential sources for deviations from the standard WLC behavior. Former studies of the influence of attractive interactions on the conformational behavior of semiflexible polymers have revealed the first insights into the phase structure [8]. Investigations of orientational order [9], nonequilibrium properties [10,11], and entanglement [12] in melts and liquid crystals of semiflexible polymers have also been the subjects of recent studies.

Recent experiments found evidence for higher flexibility in DNA at short length scales [13,14], implying the need to better understand packing properties of DNA at small length scales. Further motivation is found in recent studies of nanoscale fabrication processes [15] that utilize finite-size DNA segments to create highly ordered three-dimensional objects.

On the other end, there has also been much interest in studying the structural behavior of fully flexible polymers. This class of polymers exhibits a variety of dense structural phases governed entirely by competing volume-exclusion effects and attractive interactions between nonbonded

monomers. Interactions between pairs of monomers i and j (bonded or nonbonded) are usually modeled by the standard Lennard-Jones (LJ) potential $U_{\mathrm{LJ}}(r_{ij}) = \epsilon[(r_0/r_{ij})^{12} - 2(r_0/r_{ij})^6]$, where r_{ij} is the distance between two nonbonded monomers. The parameters r_0 and ϵ fix the length and energy scale of this interaction (we set $\epsilon = 1.0$ and $r_0 = 1.0$). The elasticity of bonds is described by the bond potential $U_{\mathrm{bond}}(r_{ii+1})$, which is a combination of the LJ potential and the finitely extensible nonlinear elastic (FENE) potential [16] $U_{\mathrm{FENE}}(r_{ii+1}) = -KR^2 \ln\{1 - (r_{ii+1}/R)^2\}^{1/2}$ (we set $R = 1.2$ and $K = 2$). The LJ parameters for the bonds were chosen to make the potential U_{bond} minimum at bond length $r = r_0 = 1$. Thus, the total energy of a given conformation reads $E_{\mathrm{flex}} = \sum_{i<j} U_{\mathrm{LJ}}(r_{ij}) + \sum_i U_{\mathrm{bond}}(r_{ii+1})$.

For this class of polymers, surface effects are essential and thus transition properties depend sensitively on the chain length [17–19]. This complex transition behavior requires sophisticated computer simulation methodologies [20], such as generalized-ensemble Monte Carlo simulation algorithms like multicanonical sampling [21] or the Wang-Landau (WL) method [22].

In this Letter, we will discuss the influence of bending rigidity on the ability of classes of single elastic polymers to form stable compact structural phases. Our goal is the construction of the entire conformational phase diagram, parametrized by temperature T and bending stiffness κ; the latter is a constant material parameter, similar to $\bar{\kappa}$ in the continuous wormlike-chain model. Therefore, we extend the model for elastic, flexible polymers by a WLC-like bending term, $E = E_{\mathrm{flex}} + E_{\mathrm{bend}}$. In our discrete polymer model, $E_{\mathrm{bend}} = \kappa \sum_l (1 - \cos\theta_l)$, where θ_l represents the angle between adjacent bonds. We focus our study on polymers with $N = 30$ monomers. Thus, the system exhibits sufficiently high cooperativity, enabling the formation of stable structural phases, although finite-size effects are not negligible. It therefore possesses essential features that are common to a large class of biomolecules. We note that the general phase structure also remains intact for larger systems; however, one has to keep in mind that the finite-size effects essentially affect the formation of structured (pseudo)phases of such small systems and that geometric assemblies of monomers depend on N [17–19]. For this purpose, we have also investigated and compared these phases for different chain lengths in the interval $13 \leq N \leq 55$.

For an efficient and systematic simulation of this model, we have developed a variant of the WL method that samples the entire (E, κ) space in a single simulation. Monte Carlo updates, including energy-dependent random displacement, reptation, and end cut-and-join moves [17], as well as κ shifts, are accepted with the probability $p((E_1, \kappa_1) \to (E_2, \kappa_2)) = \min[g(E_1, \kappa_1)/g(E_2, \kappa_2), 1]$, where $g(E, \kappa)$ is the number of states for given values of energy and stiffness. The iterative estimation of $g(E, \kappa)$ is guided by the WL procedure that has previously been used for the study of flexible polymers [17]. For redundancy checks and error estimation, a minimum of 15 independent runs were performed and averaged for each chain length.

For the identification of the structural phases, we investigate thermal fluctuations of energy and radius of gyration as functions of T and κ by canonical statistical analysis. The quantities we consider here are defined by temperature derivatives of the mean energy, in the form of the specific heat $c_V(T, \kappa) = (1/N)\partial\langle E\rangle/\partial T = (\langle E^2\rangle - \langle E\rangle^2)/Nk_{\mathrm{B}}T^2$ and of the square radius of gyration $\partial\langle R_g^2\rangle/\partial T = (\langle ER_g^2\rangle - \langle E\rangle\langle R_g^2\rangle)/k_{\mathrm{B}}T^2$ ($k_{\mathrm{B}} \equiv 1$ from now on). Both quantities can be considered as thermodynamic landscapes in parameter space (T, κ), where cooperative behavior (resulting in structural transitions) is signaled by regions of high thermal activity ("ridges"). It should be noted that this kind of comparative canonical analysis leaves a systematic uncertainty in the estimation of transition lines, yielding a complete but rather qualitative picture of the transition behavior. This is due to the finite size of the system, and, for this reason, structural phases identified here should not be confused with phase transitions in the thermodynamic sense.

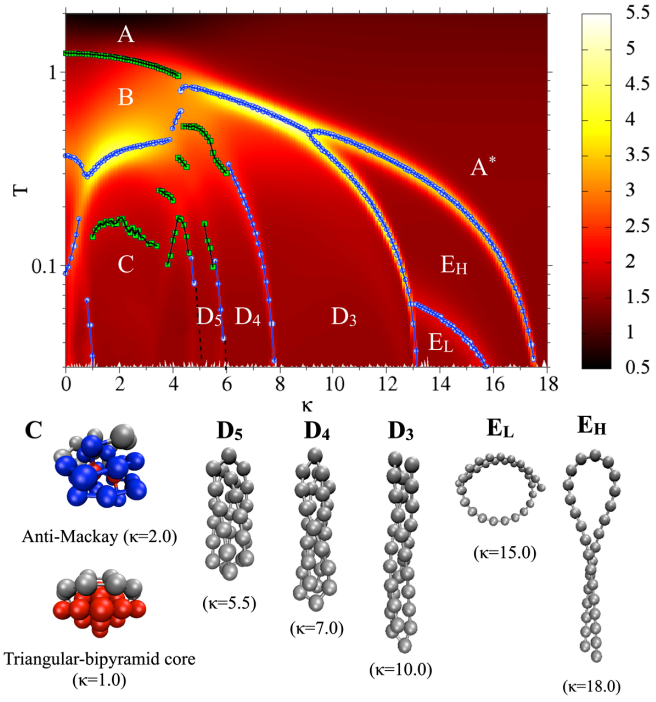

FIG. 1 (color online). Surface plot of the specific heat for classes of polymers with $N = 30$ monomers as a function of temperature T and stiffness κ. Brighter colors correspond to higher thermal activity, signaling structural transitions. For a large number of κ values, locations of peaks and shoulders are emphasized by circles and squares, respectively, for easier identification of transition points. Conformational phases are labeled as follows: A, random coil; A^*, random rodlike; B, liquid globular; C, solid globular; D_m, rodlike bundles with m segments; and E, toroidal.

Our results for the energetic and the structural representations of the conformational phase diagram are shown in Fig. 1 (specific heat) and Fig. 2 (structural fluctuations). Both landscapes feature essentially the same information, implying the robustness of our approach. Below Fig. 1, minimum-energy conformations (MECs), representative for the structured phases (C, D, and E) in different κ intervals, are depicted. Peak and shoulder positions have been identified in the specific heat and $\partial\langle R_g^2\rangle/\partial T$ landscapes. Their locations are marked by symbols in the surface maps in Figs. 1 and 2, revealing transition lines between a number of unique states as a function of T and κ. The obtained structural phase diagram can then be clearly separated into three major regions: random coils and rods (A), liquid globules (B), and a variety of structured phases (C, D, and E).

Phase A can be considered as the vapor phase, at least near the entropy-dominated flexible edge (small κ values), where conformations are unstructured, random coils. With increasing stiffness, the entropic freedom gets gradually more restricted by the competing bending energy and structural fluctuations become smaller. In the regime A^*, where bending cannot be compensated anymore by entropic freedom (as in A) or by the energetic gain of attractive nearest-neighbor interaction (cyclization in the toroidal phases E), only rodlike structures persist. Energetic and structural fluctuations indicate a transition between A and A^*.

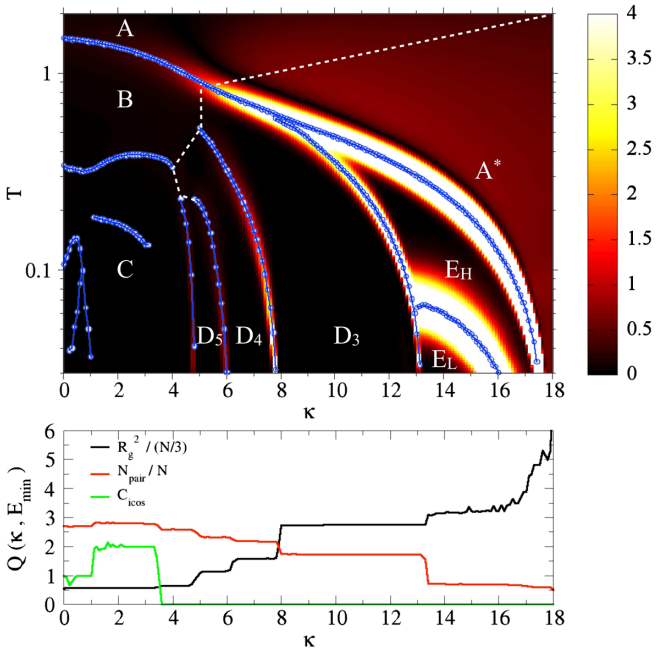

FIG. 2 (color online). Top: Surface plot of $d\langle R_g^2\rangle/dT$, with peaks indicated by circles and subsequent lines highlighting conformational phase boundaries. Dashed lines indicate transitions suggested by the Khalatur parameters and the specific heat. Bottom: Square radius of gyration R_g^2, the relative number of pairwise monomer-monomer contacts N_{pair}/N, and the number of icosahedral cores C_{icos} for the MECs in dependence of κ, providing insight into differences of structural properties in the low-T phases.

The structural behavior of the polymer in region A^* resembles that of a WLC, and the persistence length is virtually constant and identical with the bending stiffness parameter, $\xi \approx \kappa$. Note that only this single phase in the complex phase diagram accommodates the classes of polymers that are typically considered as "semiflexible" and "stiff."

Phase B is the only structural phase where the polymer behaves like a liquid. As is typical for rather small polymer chains, the Θ collapse transition between A and B is not pronounced in the specific heat (only shoulders in Fig. 1) but it is very strong in the structural response quantities (see Fig. 2). Conformations in B are unstructured but compact and globular. Beyond a threshold value ($\kappa \approx 4.5$ for $N = 30$), there is no longer a liquid phase and the polymer crystallizes abruptly out of the "vapor" phase A upon cooling (order-disorder transition).

The most intricate conformational macrostates belong to the compact solid, globular phase C. In this regime, the polymer maximizes the number of pairwise monomer-monomer contacts and is, therefore, highly energy dominated. Since monomers in the center of the globule can only possess a maximum number of 12 neighbors, structure formation is guided by a competition of surface and volume effects. Small changes in T or κ result in restructuring. Thus, C is actually a composition of several rather glasslike subphases, dominated by energetically metastable states. Close to the flexible limit ($\kappa = 0$), finite polymers tend to form structures with local symmetries. Most amazing are the perfectly icosahedral structures known for flexible polymers with "magic" chain lengths $N = 13, 55, 147, \ldots$ [17,18]. The MECs of the 30mer are found to contain one ($\kappa < 1.0$) or two ($1.0 < \kappa < 3.5$) icosahedral cores; see the bottom panel of Fig. 2.

The characteristic structures in the D_m phases are bundles of rodlike fibers, where m is the number of fibers in the bundle (the number of turns is $m - 1$). The number of turns or fibers is governed by the competition of the energetic gain in forming monomer-monomer contacts and the energetic "penalties" caused by the turns. Therefore, the precise substructure of this phase strongly depends on N. Since turns become energetically more costly for larger values of κ, the number of fibers decreases correspondingly with increasing κ, along with the number of monomer-monomer contacts. This can be seen in the bottom panel of Fig. 2, where the relative number of pairwise contacts N_{pair}/N for the MECs is plotted in dependence of κ. This quantity exhibits a noteworthy, but rather slight, steplike decrease in the D phase (note the pronounced steplike increases in the MECs' radii of gyration in this regime shown in the same figure). Therefore, the length of the bundle $N_{\text{bundle}} \approx N/m$ can easily be estimated by the

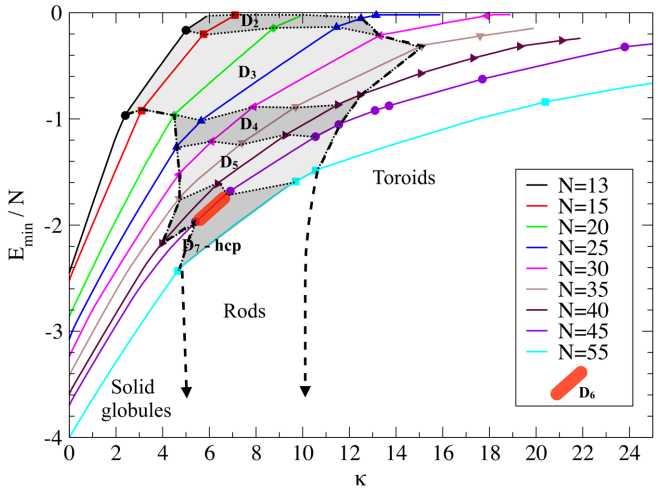

FIG. 3 (color online). MEC energies for chains with $13 \leq N \leq 55$ in dependence of κ. Symbols indicate points of qualitative structural changes of the MECs (e.g., a change from solid to a rod or a change within rods, D_3 to D_4, etc.). Dashed lines with arrowheads are visual indicators of possible behavior for $N > 55$ and are not any formal extrapolation.

ratio of the relevant competing energy scales: $N_{\text{bundle}} \approx \kappa/\epsilon$ (where $\epsilon = 1$ in our units), such that the degree m of the D phases is empirically given by $N\epsilon/\kappa$ [compare with examples of D_m conformations in the bottom diagrams of Fig. 1]. Note that, as N increases, higher-order geometries with larger contact numbers become present, e.g., cylindrical hcp-based conformations appear for $N \geq 45$, in which case the simple relationship between m and κ no longer holds (see also Fig. 3).

Eventually, we find toroidal loops in phase E_{L} and hairpins in E_{H}, very characteristic for semiflexible polymers near the WLC regime under the influence of nonbonded interactions. These structures resemble spontaneously cyclized double-stranded DNA structures that are deemed essential for gene regulation processes in cells [3,4]. The dominant structural behavior in both regions is difficult to identify, and the Khalatur parameters [8] were used to unravel structural details of E_{L} and E_{H}. Toroids found in the E_{L} region undergo subtle changes in structure, making them difficult to analyze. Changes occurring with κ are typically seen in the inner radius of the conformation; however, larger chains exhibit significant changes in the stacking of loops within a toroid. The MEC analysis in Fig. 2 does not allow for a clear distinction between these cases, even more so as structures from the boundary region of D_3 and E_{H} mix.

Supporting our conclusion that the phase structure obtained for $N = 30$ is qualitatively similar for other system sizes, Fig. 3 shows the κ dependence of MEC energies for chain lengths in the interval $13 \leq N \leq 55$. This is particularly relevant for the structural phases dominated by finite-size effects such as the D phases. D_2 through D_5 are consistent with previous discussions, but this pattern is upset for $N \geq 35$, where a jump to D_7 (hexagonal close packing) occurs, with only a single instance of D_6 (red lozenge-shaped area) occurring for $N = 45$ (within our resolution of N).

In conclusion, we have investigated the full conformational behavior of all polymer classes from flexible to stiff, i.e., the dependence on temperature and bending stiffness. For this purpose, we have specifically performed WL computer simulations of a generic energy model for flexible polymers, extended by a standard bending term. This study focused on bridging the apparent gap between recent studies on flexible polymers and the wormlike-chain model typically used to describe the behavior of semiflexible polymers. Consequently, we not only identified the general phases between these limits but also discussed the influence of finite-size effects on structure formation. This is a significant problem as it has become apparent that the structural properties of classes of short semiflexible biomolecules can significantly deviate from the standard wormlike-chain behavior. This implies that the changed structural behavior needs to be considered in the understanding of biomolecular processes on short length scales and also in the nanofabrication of molecular devices.

This project has been partially supported by NSF Grants No. DMR-0810223 and No. DMR-1207437.

*dseaton@mit.edu
†stefan.schnabel@itp.uni-leipzig.de
‡dlandau@hal.physast.uga.edu
http://www.csp.uga.edu
§bachmann@smsyslab.org
http://www.smsyslab.org

[1] O. Kratky and G. Porod, J. Colloid Sci. **4**, 35 (1949).
[2] G. Caliskan, C. Hyeon, U. Perez-Salas, R. M. Briber, S. A. Woodson, and D. Thirumalai, Phys. Rev. Lett. **95**, 268303 (2005); C. Hyeon, R. I. Dima, and D. Thirumalai, J. Chem. Phys. **125**, 194905 (2006).
[3] T. E. Cloutier and J. Widom, Mol. Cell **14**, 355 (2004).
[4] J. Yan and J. F. Marko, Phys. Rev. Lett. **93**, 108108 (2004).
[5] P. A. Wiggins, R. Phillips, and P. C. Nelson, Phys. Rev. E **71**, 021909 (2005); P. A. Wiggins and P. C. Nelson, Phys. Rev. E **73**, 031906 (2006).
[6] D. A. Sivak and P. L. Geissler, J. Chem. Phys. **136**, 045102 (2012).
[7] H.-P. Hsu, W. Paul, and K. Binder, Europhys. Lett. **92**, 28 003 (2010); H.-P. Hsu and K. Binder, J. Chem. Phys. **136**, 024901 (2012).
[8] J. A. Martemyanova, M. R. Stukan, V. A. Ivanov, M. Müller, W. Paul, and K. Binder, J. Chem. Phys. **122**, 174907 (2005).
[9] F. Affouard, M. Kröger, and S. Hess, Phys. Rev. E **54**, 5178 (1996).
[10] M. Kröger, Phys. Rep. **390**, 453 (2004).
[11] R. G. Winkler, J. Chem. Phys. **133**, 164905 (2010).

[12] R. Everaers, S. K. Sukumaran, G. S. Grest, C. Svaneborg, A. Sivasubramanian, and K. Kremer, Science **303**, 823 (2004).

[13] J. A. Abels, F. Moreno-Herrero, T. van der Heijden, C. Dekker, and N. H. Dekker, Biophys. J. **88**, 2737 (2005); P. A. Wiggins, T. van der Heijden, F. Moreno-Herrero, A. Spakowitz, R. Phillips, J. Widom, C. Dekker, and P. C. Nelson, Nat. Nanotechnol. **1**, 137 (2006).

[14] C. Yuan, H. Chen, X. W. Lou, and L. A. Archer, Phys. Rev. Lett. **100**, 018102 (2008).

[15] H. Dietz, S. M. Douglas, and W. M. Shih, Science **325**, 725 (2009).

[16] R. B. Bird, C. F. Curtiss, R. C. Armstrong, and O. Hassager, *Dynamics of Polymeric Liquids* (Wiley, New York, 1987), 2nd ed.

[17] D. T. Seaton, T. Wüst, and D. P. Landau, Comput. Phys. Commun. **180**, 587 (2009); Phys. Rev. E **81**, 011802 (2010).

[18] S. Schnabel, T. Vogel, M. Bachmann, and W. Janke, Chem. Phys. Lett. **476**, 201 (2009); S. Schnabel, M. Bachmann, and W. Janke, J. Chem. Phys. **131**, 124904 (2009).

[19] S. Schnabel, D. T. Seaton, D. P. Landau, and M. Bachmann, Phys. Rev. E **84**, 011127 (2011).

[20] S. Schnabel, W. Janke, and M. Bachmann, J. Comput. Phys. **230**, 4454 (2011).

[21] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991); Phys. Rev. Lett. **68**, 9 (1992).

[22] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001); Phys. Rev. E **64**, 056101 (2001); Comput. Phys. Commun. **147**, 570 (2002).